Preserving Data Integrity in Peer-to-Peer Systems

Mema Roussopoulos
Harvard University

What is Peer-to-Peer?

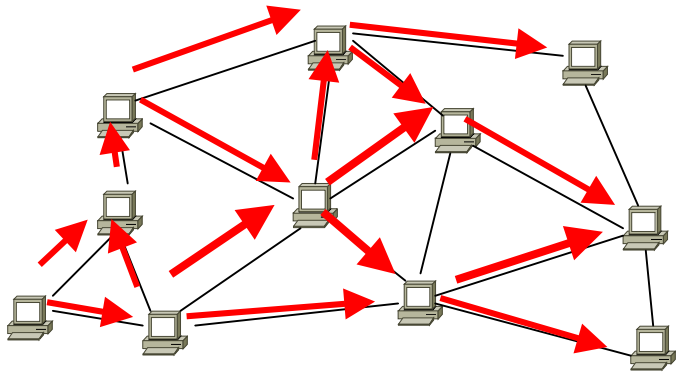
- Peer-to-Peer (P2P) concept:
  - Leverage idle resources
- Definition:
  - Self-organizing distributed system
  - Nodes provide and receive services in cooperative effort
- Features:
  - Scalability
  - Availability
  - Fault tolerance, etc.

Search in P2P

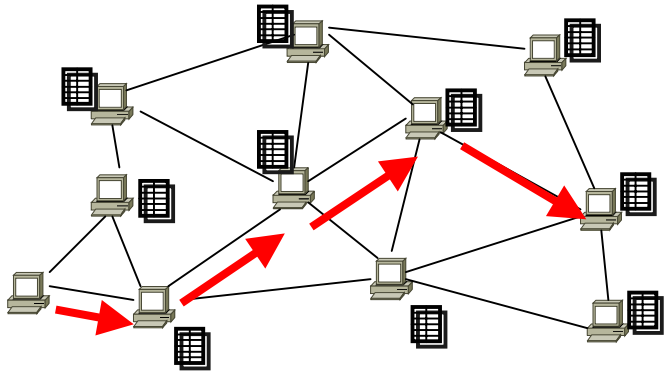
- A key operation
- Search query:
  - given name or keyword attributes of content, where is it?
- Search response:
  - a set of index entries pointing to replica nodes storing the content
- Index entry:
  - (key, value) pair
  - Key = name of content
  - Value = IP address of serving peer

Great for (illegal) file-sharing!!

- Unstructured
  - Query flooding
  - Gnutella, FreeNet



- Structured
  - Single query path
  - CAN, Chord, Pastry, Tapestry



- Anything else?

Problem Characteristics

- Participating entities are
  - Autonomous
  - Mutually distrustful
  - Mutually dependent

Example: Digital Preservation of on-line published material

Traditional Library Model

- Goal: Preserve access to important documents for posterity
- On behalf of their institution, libraries
  - Acquire and distribute lots of paper copies of important materials
  - Give access to local readers
  - Lend copies to other libraries
- It is hard to destroy all copies

Transition to Digital Media

- Resources: electronic journals, proceedings, etc.
  - Publisher rents access to materials
- Problem:
  - Libraries no longer own the bits
  - Libraries vulnerable to disappearance, failure, misbehavior of publishers

Example: Time Magazine

George Bush and Brent Scowcroft

WHY WE DIDN'T REMOVE SADDAM

Essay by Bush Sr. published in paper-version of March 2, 1998 issue

Online Version Removed

- Online version has disappeared
- Online table of contents modified
- It's as if article never existed in this issue!

LOCKSS Goals

- Lots Of Copies, Keep Stuff Safe
- Emulate traditional model for on-line publishing
- Make it easy for libraries to
  - Own, rather than lease, materials
  - Preserve and provide access to local patrons
- Make it easy for publishers to
  - Provide content for preservation and archiving
  - Without the responsibility for perpetual presence
  - With minimal risk to their business model

LOCKSS Approach

- Build p2p community of libraries
- Audit and repair their contents with
  - No centralized control (Autonomous)
  - Mutual distrust
  - Very low-cost hardware, operation and administration (Mutually Dependent)
  - A long-term horizon; I.e., preserve for decades
- Must anticipate natural bit degradation
- Must anticipate sustained attacks

Opinion Polls

- Obtaining full consensus is difficult
- Each peer holds
  - Reference list of peers it has discovered
  - Friends list of peers it knows externally
- Periodically (faster than rate of bit rot)
  - Takes a sample of the reference list
  - Invites the chosen peers to send a hash of their copy of the document

Opinion Polls (cont'd)

- Peer compares votes with local copy
- If landslide agreement, the peer is happy
- If landslide disagreement, the peer repairs
  - To repair, the peer gets the copy of somebody who disagreed and then reevaluates the same votes
- If poll is inconclusive, the peer raises alarm
  - Alarms are built-in intrusion detection

Reference List Update

- Take out voters in the poll
  - So that the next poll is based on different group
- Replenish with some “strangers” and some “friends”
  - Strangers: Accepted nominees proposed by voters
  - Friends: From the friends list
  - The measure of favoring friends is called *churn factor*

Adversary Goals

- Top adversary goal: **Stealth Modification**
  - Modify documents unobtrusively
  - Hard to reinstate original content after large proportion of peers have had their documents modified
- Other goals
  - Slow the system down
  - Discredit the system
  - Obtain benefits without contributing
  - Obtain content illicitly

LOCKSS Defenses

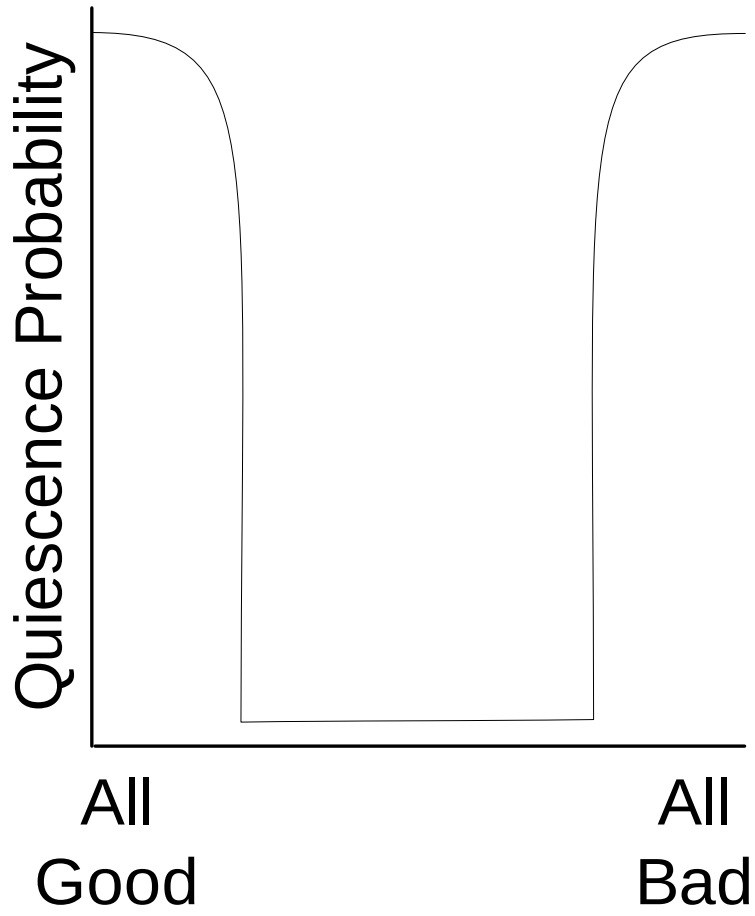
- Limit the rate of operation
- Bimodal system behavior
- Churn friends into reference list

Limit the rate of operation

- Peers determine their rate of operation autonomously
  - Adversary must wait for the next poll to attack through the protocol
- No operational path is faster than others
  - Artificially inflate “cost” of cheap operations
  - No attack can occur faster than normal ops

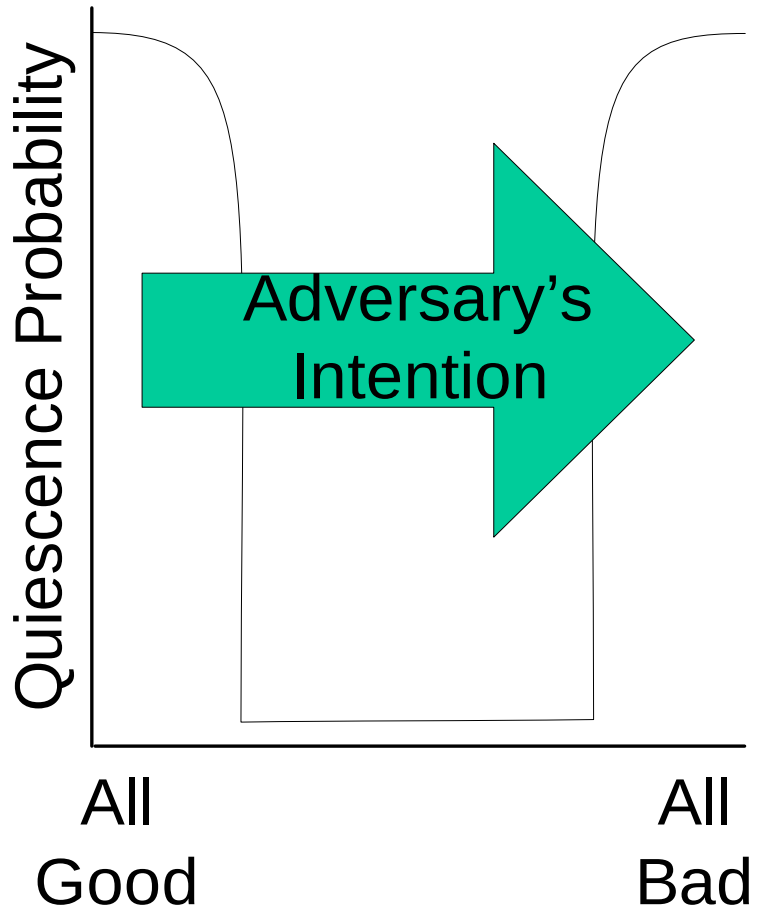
Bimodal System Behavior

- When most replicas are the same, no alarms
- In between, many alarms
- To get from mostly correct to mostly wrong replicas, system must pass through “moat” of alarming states



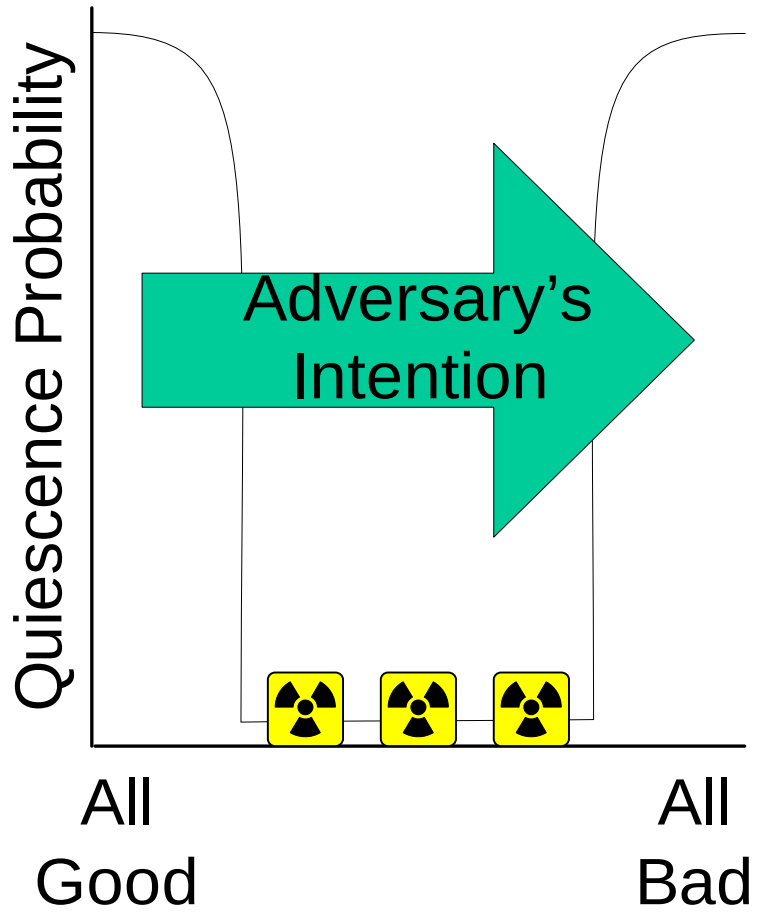
Bimodal System Behavior

- When most replicas are the same, no alarms
- In between, many alarms
- To get from mostly correct to mostly wrong replicas, system must pass through “moat” of alarming states



Bimodal System Behavior

- When most replicas are the same, no alarms
- In between, many alarms
- To get from mostly correct to mostly wrong replicas, system must pass through “moat” of alarming states



Churn Friends into Reference List

- Churn adjusts the bias in the reference list
- High churn favors friends
  - Reduces the effects of Sybil attacks
  - But offers easy targets for focused attack
- Low churn favors strangers
  - It offers Sybil attacks free reign
    - Bad peers nominate bad; good peers nominate some bad
  - Makes focused attack harder, since adversary can predict less of the poll sample
- Goal: strike a balance

Evaluation Methodology

- Model a very powerful, realistic adversary
- Identify major goals of adversary attacks
- Devise and implement rational strategies
- Measure the impact of each strategy
  - locally (on library patrons)
  - globally (on document survival)

Adversary Model

- Unlimited identities
  - Purchased (cheap) or spoofed (cheaper)
- Exploits common implementation bugs
  - Exploited peer is *subverted*
- Perfect coordination
  - Instantaneous communication with and control of subverted peers
  - Load balancing of attack effort
  - Flawless content preservation

Stealth Modification Strategy

- A peer's reference list affects outcomes of polls it will call
- The stealth adversary
  - First, quietly gains a strong foothold in the reference list of a peer
  - Then, attacks when a poll will be landslide win in favor of adversary's copy
  - Must consistently win polls to succeed

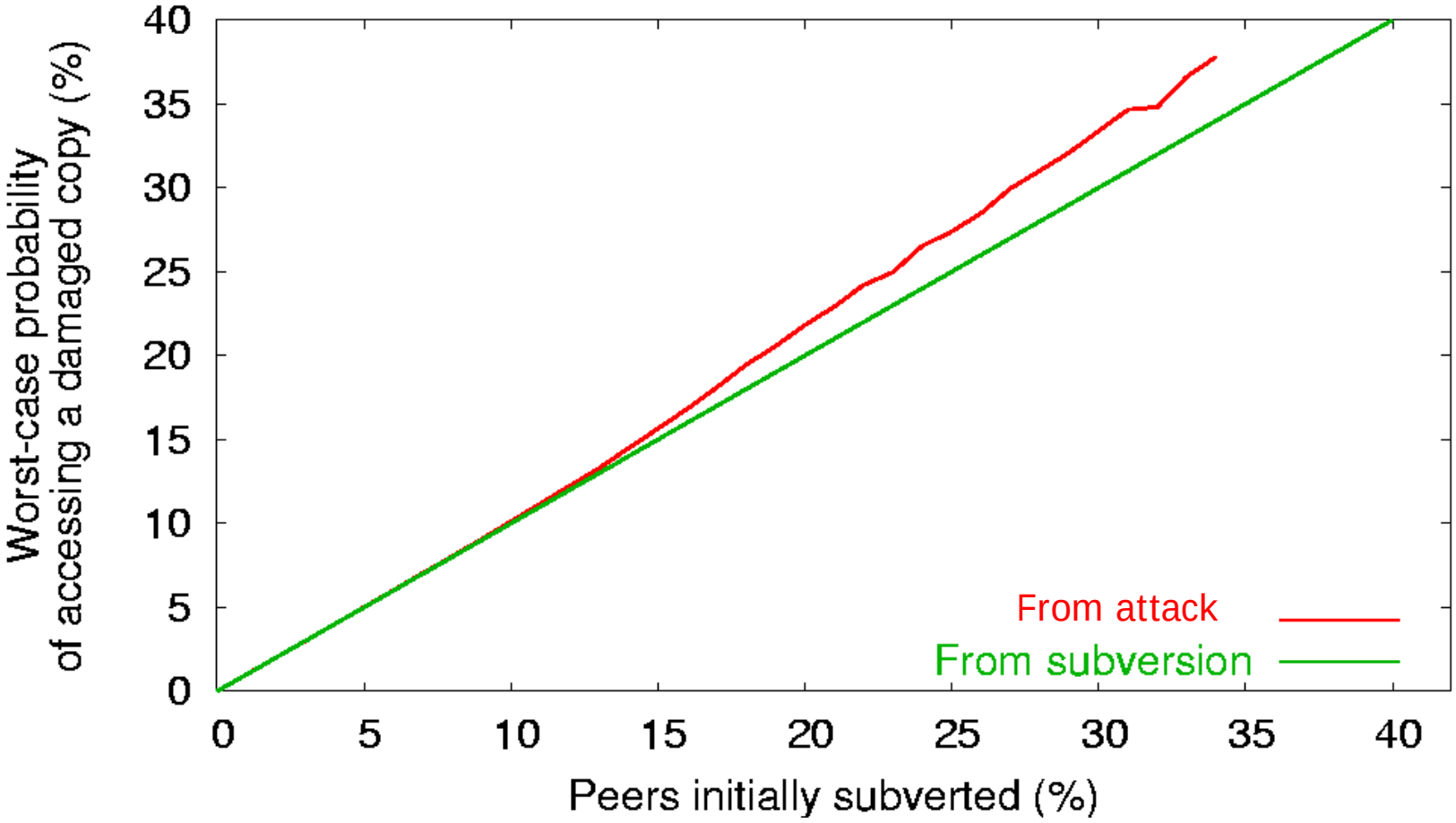
Evaluation

- We use Narses, an application-layer protocol simulator
- Scenarios
  - 1000 original peers, in clusters of friends
  - Initially, 0 – 40% are subverted
  - Lurk for up to 20 years
  - Attack for up to 10 more years
  - Report worst-cases over ~200 runs per data point (recent results)

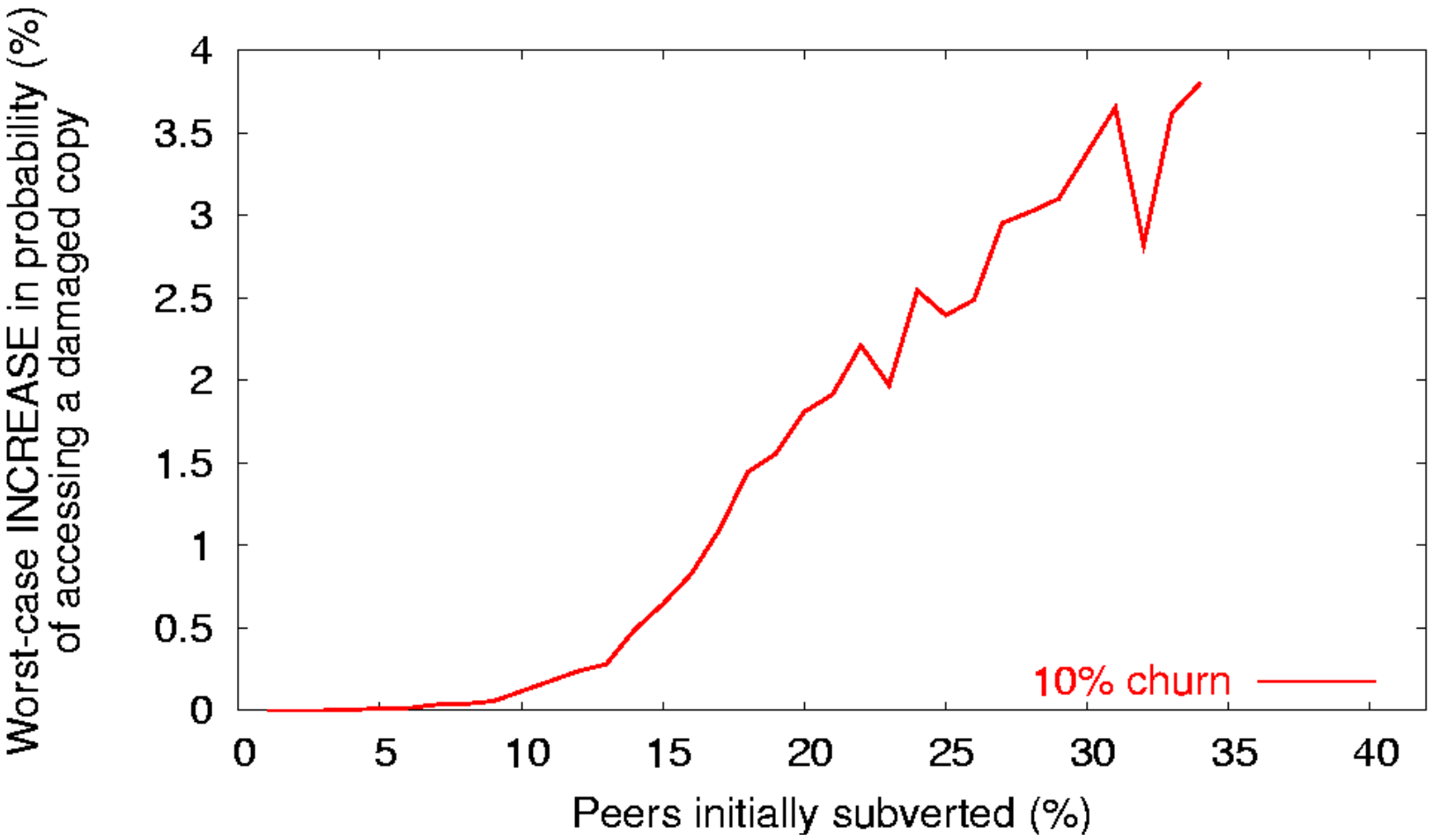
Metrics

- Metrics
  - What's the probability that an access reaches a bad replica
  - What's the probability that the document is damaged irrecoverably
- **How big is the effect of the worst protocol attack on top of the effect of the initial subversion?**

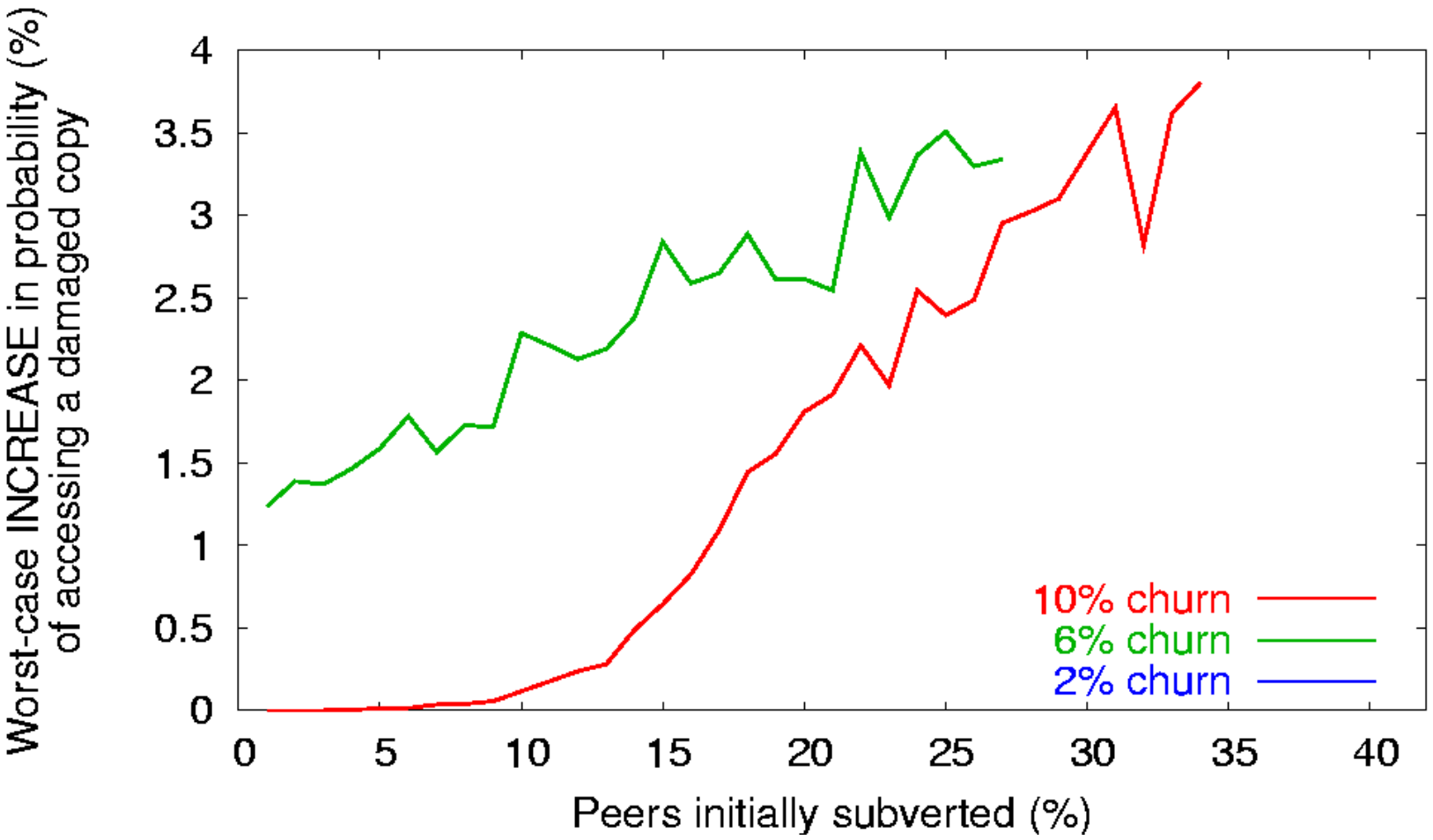
Probability of Accessing Bad Copy



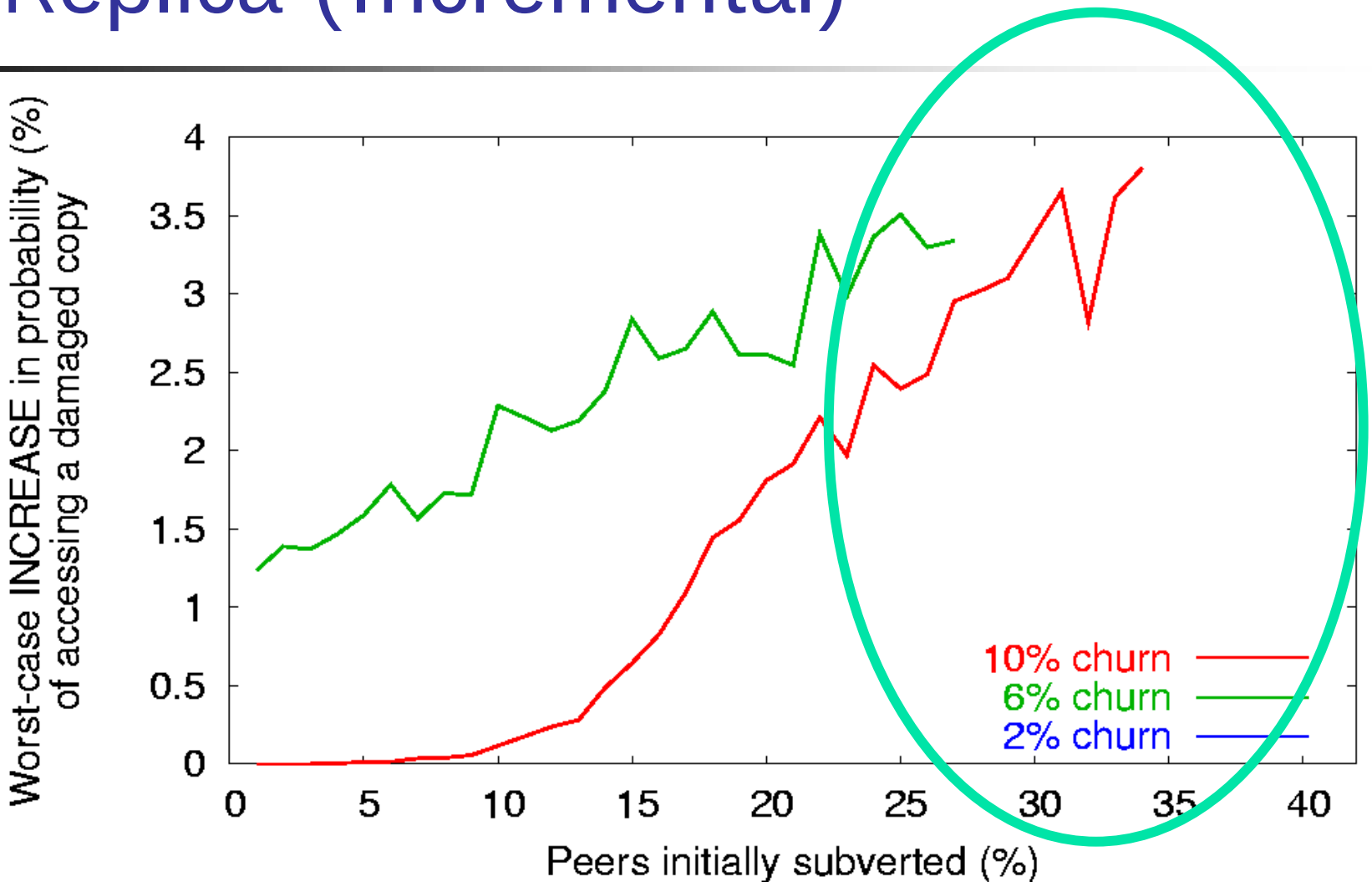
Probability of Accessing Bad Replica (Incremental)



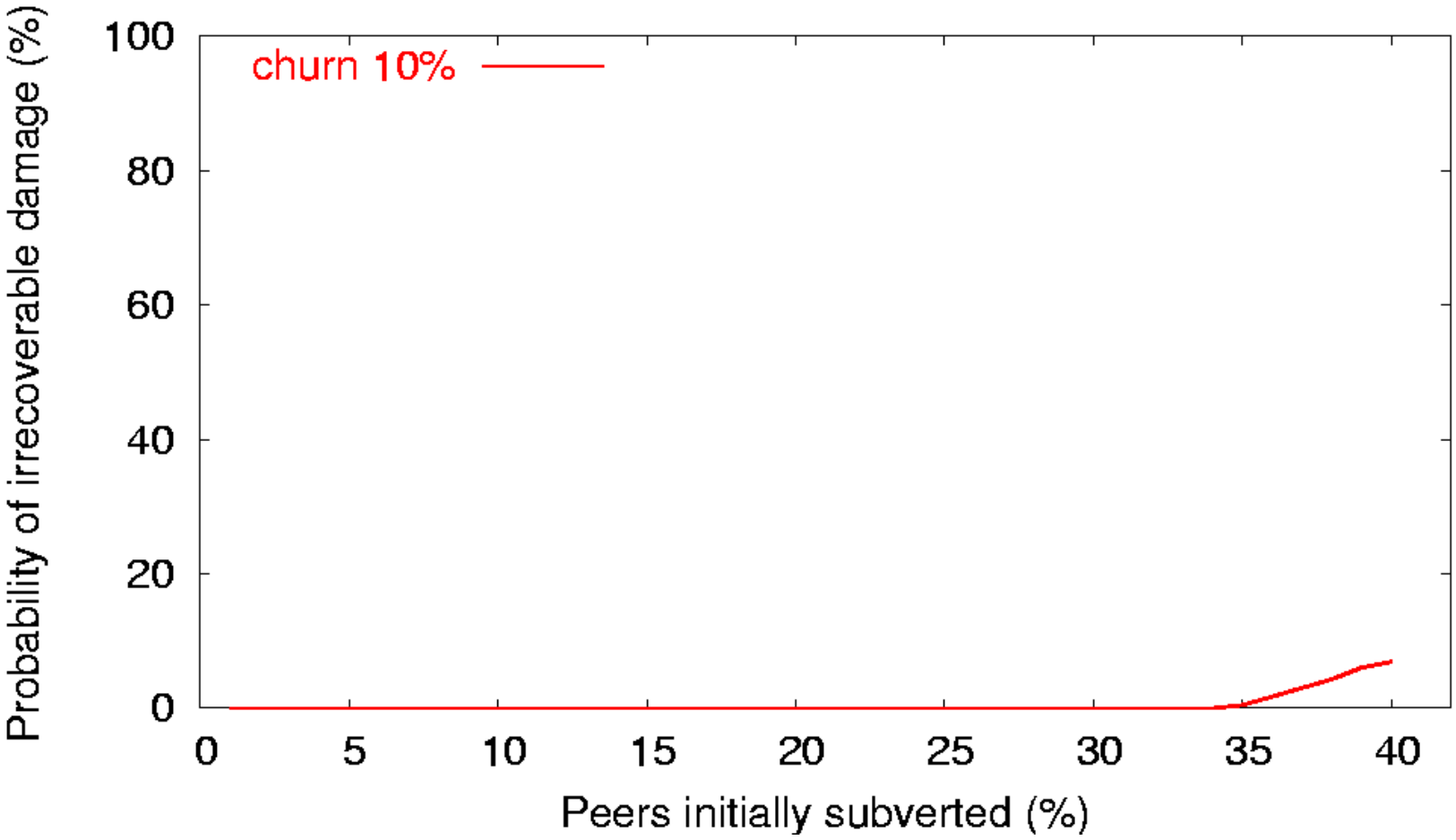
Probability of Accessing Bad Replica (Incremental)



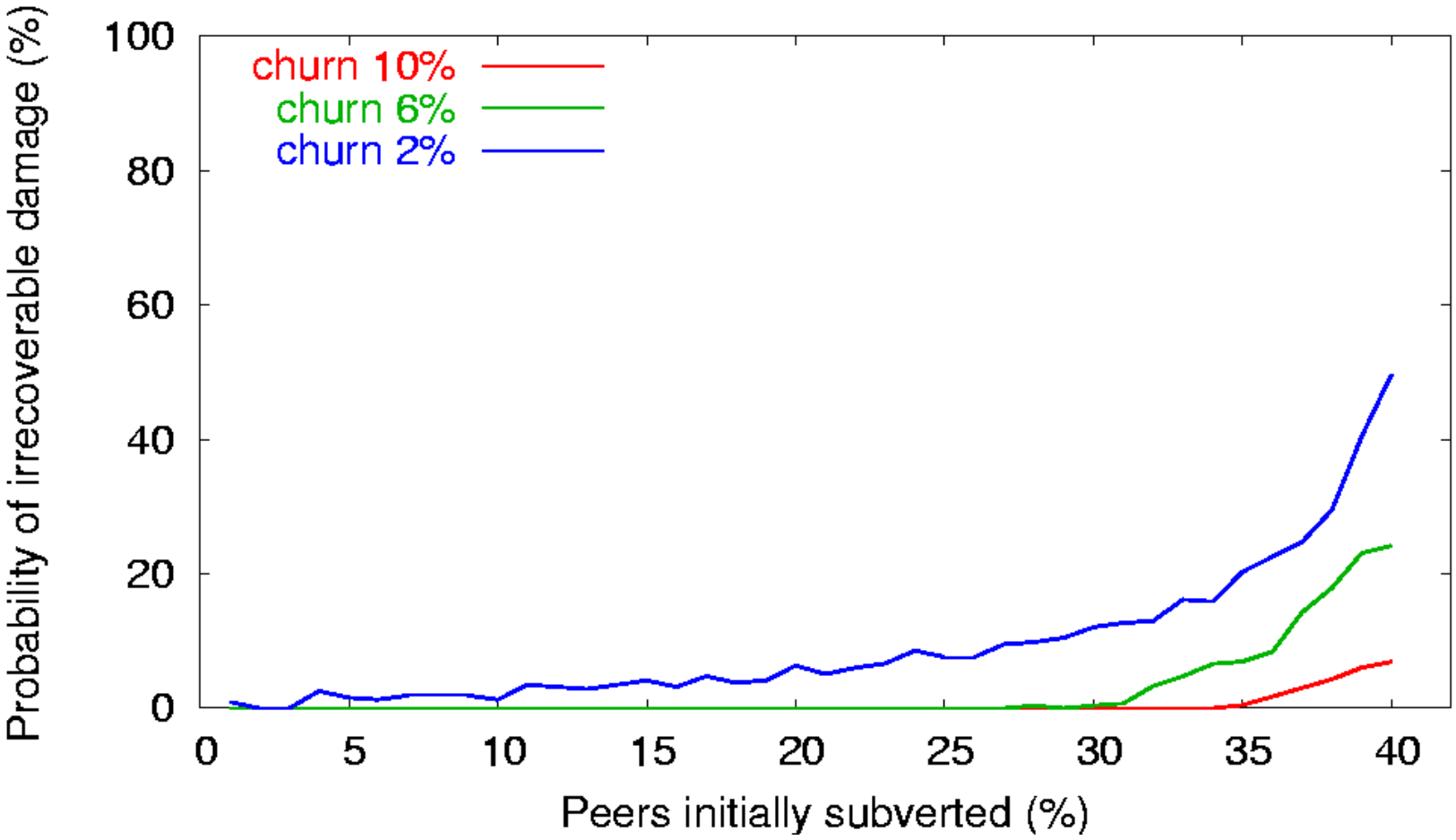
Probability of Accessing Bad Replica (Incremental)



Probability of Irrecoverable Damage



Probability of Irrecoverable Damage



Alternatives

- Use super-fabulous RAID
  - Can be complementary, but alone cannot ensure survivability when failures do occur (e.g., human error)
- Encrypt or sign to ensure integrity
  - Preserving public keys just as hard a problem
- Boost efficiency with erasure codes etc.
  - Storage space is not an issue
  - All replicas must be whole

Next Steps

- Explore the parameter space
  - What quorum sizes are necessary?
  - Frequency of polls vs. rate of undetected medium faults vs. probability of adversary success
- Enlarge bestiary of attackers
  - Attrition attacks (e.g., DDoS)
  - Hybrid attacks (e.g., stealth modification during DDoS weakening)
- Expand to other application domains

Applications

- Academic journals
  - Append-only updates
- Government documents
  - Large number, frequent updates
- High-resolution scans of artwork
  - “Rare-bits”
- Scientific data
  - Large volumes (terabytes) of data

Conclusions

- P2P is more than file-sharing
- P2P good for applications with:
  - Autonomous entities
  - Mutually distrustful entities
  - Mutually dependent entities
- One example: LOCKSS, a P2P digital preservation system

Status

- Results for stealth adversary
  - Resistant to attacks for low subversions
  - Degrades gracefully for greater subversions
- Status
  - Promising results for other attacks (DDOS)
  - To be deployed at ~100 libraries across the globe in 2004
  - For more info:

http://www.eecs.harvard.edu/~mema/